

# 基于传播免疫的复杂网络可控性研究\*

吕天阳<sup>1)2)†</sup> 朴秀峰<sup>1)</sup> 谢文艳<sup>1)</sup> 黄少滨<sup>1)</sup>

1) (哈尔滨工程大学计算机科学与技术学院, 哈尔滨 150001)

2) (清华大学计算机科学与技术系, 北京 100084)

(2011年11月16日收到; 2012年2月20日收到修改稿)

复杂网络控制反映了人类对复杂系统的认识深度和改造能力. 最新研究成果基于线性系统控制理论建立了复杂网络可控性的理论架构, 能够发现任意拓扑结构的线性时不变复杂网络中控制全部节点状态的最小驱动节点集, 但是该模型未考虑免疫节点或失效节点对控制信号传播的阻断. 在继承该模型优点的前提下, 重新构建了基于传播免疫的复杂网络控制模型. 在采用分属于随机免疫和目标免疫两种策略的4个方法确定免疫节点的情况下, 分析14个真实网络的可控性. 结果表明: 如果将网络中度数、介数和紧密度指标较高的节点作为免疫节点, 将极大地提高控制复杂网络的难度. 从而在一定程度上丰富了以往模型的结论.

**关键词:** 复杂网络控制, 传播免疫, 免疫节点

**PACS:** 05.65.+b, 02.30.Yy

## 1 引言

当前, 复杂网络理论已经成为一个重要的研究范式, 用于解释互联网、社会学、生物学等诸多领域所面临的问题. 对复杂网络演化模型的探索反映了人类认识复杂系统的深度, 对复杂网络控制的研究则反映了人类改造复杂系统的能力. 复杂网络控制技术可以用于改造复杂网络以体现特定性能, 例如提高特定网络的鲁棒性和稳定性, 因此近年来逐渐兴起.

一般而言, 复杂网络控制的目标是: 对若干选取的控制节点输入信号使得复杂网络达到预期的状态. 其中, 复杂网络控制的可行性和有效性是其基础课题. 可行性是指复杂网络是否可控, 有效性是指降低控制的代价<sup>[1]</sup>. 虽然, 多数学者认为, 复杂网络的控制应当较为系统的引入控制论研究成果, 但是基于控制论思想建立复杂网络控制的理论架

构并不是一个简单的问题.

为此, 2011年5月Liu, Barabasi等基于线性系统控制理论建立了复杂网络可控性的理论模型<sup>[2]</sup> (简称LB模型). LB模型能够发现任意拓扑结构的线性时不变复杂网络的最小驱动节点集. 利用LB模型分析真实网络和若干复杂网络模型的结果表明: 1) 最小驱动节点集的大小很大程度上与度分布相关; 2) 驱动节点趋向于避开高度节点. 结论2)有别于人们的常识推断和以往的研究结论<sup>[3]</sup>, 即网络中的最具影响力的节点应该为高度节点, 控制这些节点对影响整个网络的贡献最大. 当然, 也有研究成果指出hub节点并非是在传播中最具影响力的节点<sup>[4]</sup>.

但是LB模型仅针对线性时不变复杂网络, 而且对于控制信号的传播假定过于简单. 控制信号的传播是实现复杂网络控制的重要基础. 对此, LB模型假设: 任意节点收到控制信号必然转发给所链接的节点, 即“接收即转发”. 真实网络中却普遍存在

\* 国家自然科学基金 (批准号: 60903080, 60093009), 国家科技支撑计划 (批准号: 2009BAH42B02, 2012BAH08B02), 中央高校基本科研业务费专项资金 (批准号: HEUCF100603, HEUCFZ1212, HEUCFT1206) 和黑龙江省教育厅科学技术研究 (批准号: 12513050) 资助的课题.

† E-mail: raynor1979@163.com

阻断网络传播的节点,例如疾病传播网络中的免疫节点<sup>[5]</sup>、电力网络中的失效节点、行为传播中的阻断现象<sup>[6]</sup>,后文将此类节点统称为免疫节点.

为此,本文着重研究免疫节点的出现对复杂网络可控性的影响,重新构建了基于传播免疫的复杂网络控制模型.由于新模型仍然满足线性系统可控理论的基础假设,因此继承了LB模型的优点,同时能够分析存在免疫节点的复杂网络的可控性.在采用随机免疫和目标免疫策略确定免疫节点的情况下<sup>[7]</sup>,采用新模型分析控制真实复杂网络的可行性和有效性.其中,目标免疫采用度数、介数和紧密度三个指标确定免疫节点.考察不同类型免疫节点对复杂网络控制有效性的影响.结果表明:相比于随机免疫,如果按前述三个指标识别节点的重要性,并将较重要的节点作为免疫节点,将极大的提高控制复杂网络的难度.这一结果表明:即使驱动节点倾向于低度节点,通过高度、高介数或高紧密度的节点阻断信号传播却能有效地影响网络的可控性,从而在一定程度上丰富了LB模型的第二结论.

## 2 基于线性系统控制的复杂网络可控性

本节着重介绍LB模型的基本思想,并分析其若干缺欠.首先给出控制节点、输入节点、驱动节点与最小驱动节点集的说明.若要完全控制一个复杂网络中所有节点的状态,需要施加控制信号的节点集合称为控制节点集合,用于对复杂网络中的控制节点施加输入控制信号的附加节点称为输入节点,不共享相同输入节点的控制节点集合称为驱动节点集合,足以完全控制网络中全部节点状态所需的最少的驱动节点集合称为最小驱动节点集合.

为确定驱动节点,文献[2]将线性时不变系统的可控性理论应用于复杂网络控制中,应用Lin结构可控论<sup>[8]</sup>和最小输入原理进行分析,将网络中驱动节点的求解问题转化为求解有向图中最大匹配边的问题.以下对这些理论进行简单介绍.

### 2.1 线性时不变系统的可控性

对于线性时不变系统( $A$ 表示系统中的各个对

象的状态参数, $B$ 表示控制量参数):

$$x'(t) = Ax(t) + Bu(t), \quad (1)$$

对初始时刻 $t_0(t_0 \in T, T$ 为时间定义域)和初始状态 $x(t_0)$ ,如系统满足如下条件:存在另一有限时刻 $t_1(t_1 > t_0, t_1 \in T)$ ,可以找到一个控制量 $u(t)$ ,能在有限时间 $[t_0, t_1]$ 内把系统状态从初始状态 $x(t_0)$ 控制到原点,即 $x(t_1) = 0$ ,则称 $t_0$ 时刻的状态 $x(t_0)$ 能控.若对 $t_0$ 时刻的状态空间中的所有状态都能控,则称系统在 $t_0$ 时刻状态完全能控.

根据卡尔曼能控性判据,当结构矩阵 $C = (B, AB, A^2B, \dots, A^{N-1}B)$ 满秩时,即 $\text{rank}(C) = N$ 时,系统可控.

据此建立复杂网络控制的理论架构,则对于复杂网络 $G, N \times N$ 矩阵 $A = \{a_{ij}|i, j \in [1, N]\}$ 表示网络中节点的邻接关系,其中 $a_{ij}$ 表示节点 $j$ 指向节点 $i$ 的边的权值.向量 $x(t) = (x_1(t), x_2(t), \dots, x_N(t))^T$ 表示 $N$ 个节点在时刻 $t$ 的状态, $N \times M$ 矩阵 $B = \{b_{ij}|i \in [1, N], j \in [1, M]\}$ 表示节点与外部控制信号的连接关系, $u(t) = (u_1(t), u_2(t), \dots, u_M(t))^T$ 为 $M$ 个输入节点在 $t$ 时刻的状态,其中 $b_{ij} = 1$ 表示信号 $u_j(t)$ 施加到节点 $i$ 上, $b_{ij} = 0$ 表示信号 $u_j(t)$ 未施加于节点 $i$ .因文献[2]的状态参数矩阵应用的是表示网络节点间邻接关系及关系强度的邻接矩阵,因此后文也不再区分系统的状态参数矩阵和复杂网络的邻接关系矩阵.

### 2.2 结构可控性求解方法

Lin结构可控理论指出以下三个条件等价:1)线性系统结构可控;2) $G(A, B)$ 不含不可达(inaccessibility)节点且不含扩张(dilation);3) $G(A, B)$ 由cacti结构生成.其中cacti结构为即不含不可达节点也不含扩张的最小结构,即若去除任一条边将导致不可达节点或扩张的出现.不可达,扩张和cacti结构的具体定义参见文献[2],这里不再说明.

文献[2]中结合最小输入原理证明:结构可控性问题可以通过求解有向图最大匹配问题来解决.将有向图 $G(A)$ 转换为如下二分图 $H(A)$ : $H(A) = (V_A^+, V_A^-, \Gamma)$ ,其中 $V_A^+ = \{x_1^+, \dots, x_N^+\}$ , $V_A^- = \{x_1^-, \dots, x_N^-\}$ ,分别表示状态矩阵 $A$ 的各列与各行的节点集合, $\Gamma = \{(x_j^-, x_i^-)|a_{ij} \neq 0\}$ 表示

边集. 根据二分图匹配算法求得匹配节点与非匹配节点, 其中最大匹配边指向的节点为匹配节点, 无最大匹配边指向的节点为非匹配节点.

则控制一个网络  $G(A)$ , 只需要将控制信号输入非匹配节点并且保证控制信号可达任意匹配节点即可. 求解二分图的最大匹配算法, 其算法的复杂度为  $O(N^{1/2}L)$ <sup>[9]</sup>, 其中  $N$  为网络中节点数,  $L$  为边数. 对于非匹配节点数不为 0 的网络, 驱动节点数即为非匹配节点数, 且非匹配节点即为驱动节点的一种情况. 对于非匹配节点数为 0 的网络, 驱动

节点数为 1, 网络中任一节点均可定义为驱动节点.

图 1 以一个小网络为例描述求取驱动节点的过程. 考察图 1(a) 所示拓扑结构的复杂网络, 将网络转化为图 1(b) 所示二部图结构, 考察此二部图的最大匹配边, 最大匹配边集的一种组合为图 1(c) 中红色的边集, 可见  $v_2, v_3, v_4, v_5$  为匹配节点, 故图 1(a) 的驱动节点为未匹配节点  $v_1$ , 图 1(d) 中由  $v_2 - v_3 - v_4$  形成的匹配环与  $v_5$  可以形成 cacti 结构, 即保证了各个节点的可达性, 对节点  $v_1$  的控制就可以控制整个网络全部节点的状态.

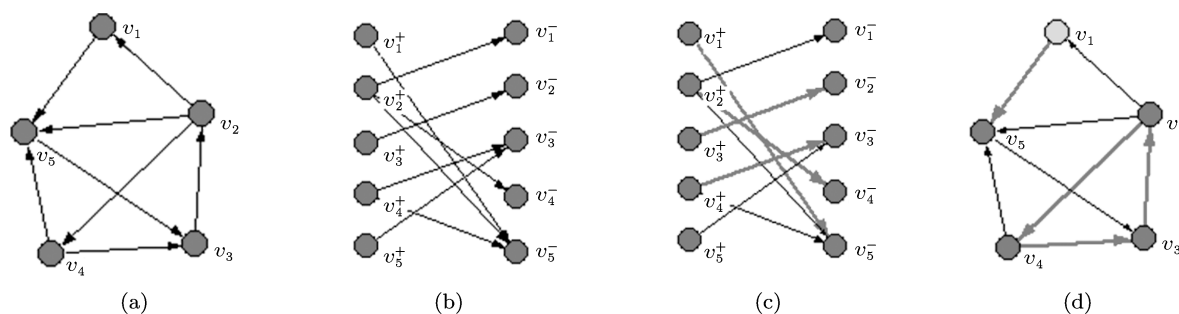


图 1 描述算法过程的例子 (a) 网络拓扑图; (b) 网络对应的二部图; (c) 二部图的最大匹配边集; (d) 网络中驱动节点分布

### 2.3 分析与讨论

在文献 [2] 发表同时出现的对 LB 模型的评论所指出的那样: 控制论和复杂网络研究两个领域顶级专家的合作, 较为成功的将控制论用于建立复杂网络可控性的理论模型 LB 模型<sup>[10]</sup>. LB 模型所给出的一些初步结论对更深入的理解复杂网络控制具有重要意义<sup>[11]</sup>, 例如“驱动节点趋向于避开高度节点”.

但是, 作为一项开拓性的研究成果, LB 模型仍然存在一些有待深入研究之处. 首先, LB 模型作为一种通用的理论框架, 必然忽视了不同类型复杂网络所面临的领域特点, 因此其结论需要一定的限定才可以推广到具体领域. 其次, LB 模型针对线性时不变复杂网络, 但是真实的复杂网络通常为非线性、时变网络. 例如, 互联网的拓扑结构的时变, 行为传播中的非线性. 再次, LB 模型对于控制信号的传播假定过于简单, 仅设定: 任意节点收到控制信号必然转发给所链接的节点, 即“接收即转发”. 但是, 真实网络中普遍存在阻断网络传播的节点, 例如行为传播研究中争鸣的两种传播模式都认为存

在阻断传播的情况<sup>[6]</sup>、以及疾病传播网络中的免疫节点、电力网络中的失效节点等.

第 3 节将基于 LB 模型针对存在免疫节点的复杂网络的可控性展开研究.

### 3 基于传播免疫的复杂网络控制

本节针对传播免疫现象, 在复杂网络控制中引入免疫节点的概念, 进一步完善复杂网络控制模型 LB 模型.

虽然复杂网络中传播的阻断现象非常常见, 但是确定实际网络中的免疫节点并非是一个简单问题, 比如人际交互网络中人的思维是无法精确预测的、也不能事先确定电力系统中的失效节点. 对此, 常见的手段是假设节点以一定的概率转发信号, 这也是病毒传播等领域常见的处理手段. 但是, 此时再应用第 2 节中的复杂网络控制理论, 则不再符合线性系统的条件. 因为当对节点施加信号输入时, 由于信号是以一定概率被转发的, 故存在不确定性, 假设节点  $n_i$  将信号以概率  $p$  转发给节点  $n_j$ , 信号进行了两次转发, 第一次转

发后节点  $n_j$  状态改变量为  $x_1(n_j)$ , 第二次转发后节点状态改变量为  $x_2(n_j)$ , 如果两次都转发成功,  $n_j$  状态的改变量为  $x(n_j)$ , 则对于线性系统应满足  $x(n_j) = x_1(n_j) + x_2(n_j)$ , 但是由于概率因素的存在, 这个等式不一定成立, 这时不满足线性系统的可叠加性. 故需要采取有别于概率的方法对这种情况重新建模分析.

本文采用去除特定节点出边的方法阻断网络中的信号传播, 从而避免基于概率的方法的缺欠. 首先给出两个定义.

**定义 1** 若信号对于网络中任意节点接收即可转发, 则称为控制信号, 记为  $u^+(t) = (u_1^+(t), u_2^+(t), \dots, u_M^+(t))^T$ .

其中  $u_1^+(t), u_2^+(t), \dots, u_M^+(t)$  为  $M$  个控制信号. 输入控制信号的节点称为控制节点, 不共享相同输入节点的控制节点称为驱动节点, 记为  $v_D$ , 驱动节点集合记为  $V_D$ .

本文通过输入免疫信号确定那些对信号起阻断作用的节点.

**定义 2** 对网络中控制信号的传播起到阻断作用的信号, 称为免疫信号, 记为  $u^-(t) = (u_1^-(t), u_2^-(t), \dots, u_M^-(t))^T$ .

其中  $u_1^-(t), u_2^-(t), \dots, u_M^-(t)$  为不同的免疫信号. 输入免疫信号的节点称为免疫节点, 记为  $\overline{v_D}$ , 免疫节点集合记为  $\overline{V_D}$ .

可见, 本文将信号分为控制信号和免疫信号两种, 控制信号的作用与 LB 模型中信号的作用一致, 而免疫信号的作用为阻断控制信号在免疫节点上的传播. 免疫信号的物理意义体现为: 中断或抵消控制信号通过此节点对网络中其他节点的状态产生的影响. 例如, 在疾病传播网络中, 免疫信号指向的节点将不再是传染源; 在电力网络中, 对节点施加免疫信号, 可以中断电信号通过此节点向其下级节点传播; 在行为传播中, 行为传播将止于施加了控制信号的节点. 在完善 LB 模型之前, 先给出有关控制信号与免疫信号的几个符号的表示, 见表 1.

表 1 信号及状态的符号表示

	输入信号	初始状态	最终状态	控制矩阵
控制信号	$u^+(t)$	$x_0^+(t)$	$x^+(t)$	$B^+$
免疫信号	$u^-(t)$	$x_0^-(t)$	$x^-(t)$	$B^-$

首先, 考察引入免疫信号的复杂网络控制系统

是否满足线性系统的条件. 由于对复杂网络施加免疫信号, 将相当于去除免疫节点的出边, 则这类节点必然不可再对控制信号进行传播, 即排除了概率因素的影响, 保证了控制信号在复杂网络中传播的确定性, 故此复杂网络仍是线性时不变系统, 其线性关系可表示为

$$(c_1 u^+(t) + c_2 u^-(t), c_1 x_0^+(t) + c_2 x_0^-(t)) \rightarrow c_1 x^+(t) c_2 x^-(t), \quad (2)$$

其中  $c_1, c_2$  为常量.

其次, 应用线性控制系统理论对复杂网络的可控性进行分析, 免疫信号的作用体现为邻接矩阵  $A$  的权值的改变, 输入免疫信号后的新邻接矩阵记为  $\overline{A}$ ,  $\overline{A} = \{\overline{a_{ij}}\}$ , 对于  $\forall i, j \in [1, N]$ ,  $\overline{a_{ij}}$  满足

$$\overline{a_{ij}} = \begin{cases} 0, & j \in \overline{V_D}, \\ a_{ij}, & \text{其他}. \end{cases} \quad (3)$$

由于免疫信号仅仅起到信号阻断的作用, 并不在网络中进行传播, 故 (1) 式中的控制矩阵  $B$  可用  $B^+$  替代, 因此, 基于传播免疫的线性控制系统的模型可表示为

$$x'(t) = \overline{A}(t) + B^+ u^+(t). \quad (4)$$

至此, 引入免疫节点的复杂网络控制模型即转化为与 LB 模型类似的模型, 其区别在  $A$  与  $\overline{A}$  差异性, 因此, 在对免疫节点加入免疫信号后, 可以应用 LB 模型求解, 免疫信号不仅起到了阻断网络中信号传播的作用, 其理论也弥补了概率转发存在的欠缺. 下面以一个小网络为例, 说明上述算法的基本过程.

对图 2(a) 所示的网络, 考察引入免疫节点后的驱动节点的分布情况. 图 2(b) 给出应用 LB 模型计算的驱动节点情况, 驱动节点为图中的绿色节点, 即  $v_6, v_7, v_9$ . 图 2(c) 给出对图 2(a) 中网络随机定义 20% 的节点 (即 3 个黄色的节点,  $v_3, v_4, v_{11}$ ) 为免疫节点后网络的新拓扑结构, 图中蓝色边为去除的出边, 对这 3 个节点施加免疫信号, 应用新模型求解驱动节点的结果为图 2(d) 中绿色节点, 即  $v_5, v_6, v_7, v_9$ . 即当网络中  $v_3, v_4, v_{11}$  接收信号但不转发时, 需要增加驱动 (控制) 节点  $v_5$ .

需要说明的是, 基于传播免疫的复杂网络控制与疾病的传播免疫的研究目的存在区别. 复杂网络控制的目标为控制全部节点的状态, 其基础的网

络模型为线性时不变网络, 添加免疫信号的目的是提高控制网络中全部节点的难度. 而疾病传播免

疫<sup>[12]</sup> 目标在于降低疾病的传播速度, 并不要求控制全部节点<sup>[13]</sup>.

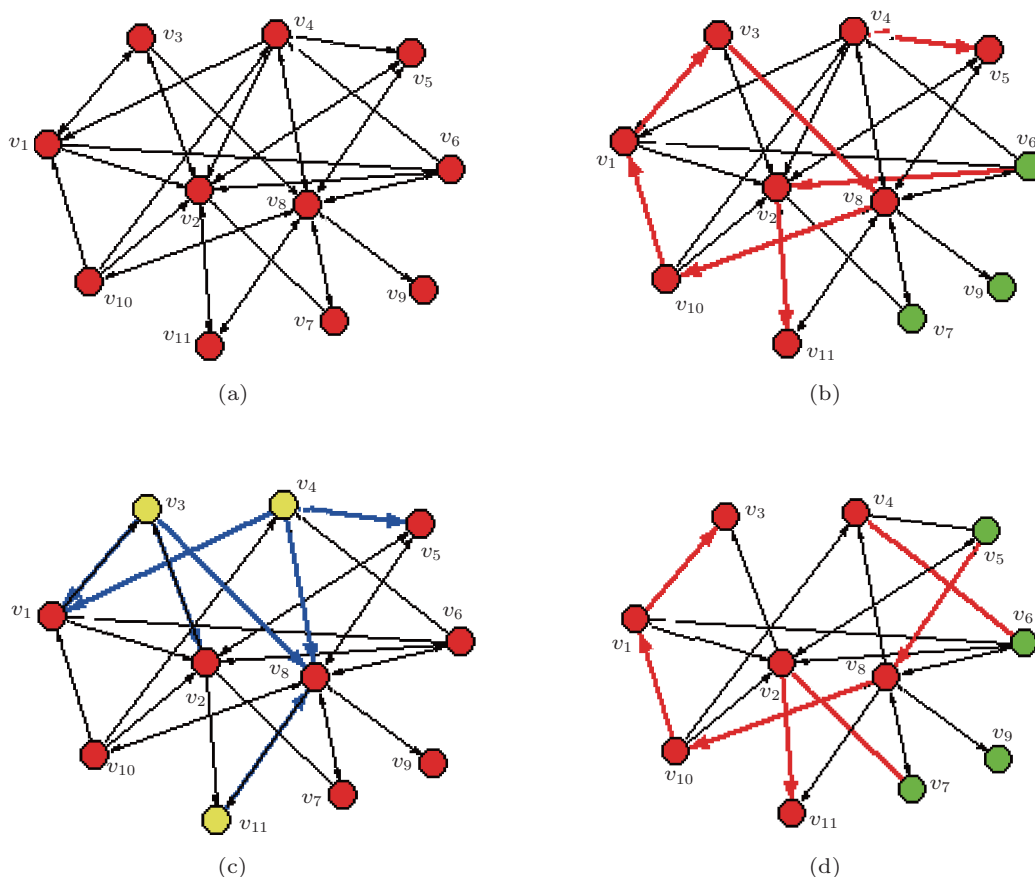


图2 网络中出现免疫节点后驱动节点的变化情况 (a) 原网络; (b) 网络中驱动节点分布; (c) 定义免疫节点后的网络拓扑结构; (d) 添加免疫节点后的驱动节点分布

## 4 实验与分析

判别对哪种类型的节点施加免疫信号可以提高复杂网络的控制难度, 这对网络控制的可行性和有效性研究具有重要意义, 本节将对这个问题进行分析.

本文采用分属随机免疫和目标免疫两种策略的四个方法确定免疫节点, 对 14 个真实网络的可控性进行分析. 对于随机免疫, 完全随机的定义网络中的免疫节点即可; 对于目标免疫, 假定网络中的重要节点为免疫节点. 通过上述两种策略对复杂网络可控性的影响判断: 节点的重要性与复杂网络可控性的关系. 对于节点重要性的度量, 本文采用复杂网络理论中考察节点中心性的指标来度量, 即度数指标、介数指标、紧密度指标, 其中, 度数指

标描述节点与周围节点产生的直接影响力<sup>[14]</sup>, 介数指标刻画网络中节点对于信息流动的影响力<sup>[15]</sup>, 紧密度指标刻画网络中节点通过网络到达网络中其他节点的难易程度<sup>[16]</sup>.

实验结果表明用这 3 个指标表示节点的重要性时, 将重要的节点作为免疫节点, 极大的增加了原网络的控制难度. 而在选取相同数量的免疫节点的前提下, 目标免疫相比随机免疫能够更有效的影响复杂网络的可控性, 提高网络的控制难度.

### 4.1 数据集分析

本文采用表 2 中 14 个数据集进行实验分析, 数据集涉及信任网络、食物网络、电力网络、神经网络、引用网络、万维网、因特网、社会交际网络和社会组织网络等多个领域.

表 2 数据集信息及驱动节点数

数据集分类	数据集	节点数	边数	平均度	驱动节点数	驱动节点比例
Trust	Wiki-Vote	7, 115	103, 689	29. 15	4736	0. 67
Food Web	Florida	128	2106	32. 91	30	0. 23
	Mondego	46	400	17. 39	19	0. 41
Power Grid	USpowerGrid	4, 941	13188	10. 68	575	0. 12
Neuronal	celegansneural	297	1017	6. 86	109	0. 37
	Caenorhabditis elegans	306	2345	15. 33	58	0. 19
Citation	hep-th	27, 770	352, 807	25. 41	5994	0. 22
WWW	nd. edu	325, 729	1,497,134	9. 19	220552	0. 68
	polblogs	1, 224	19, 025	31. 19	418	0. 34
Internet	p2p-2	8, 846	31, 839	7. 20	5111	0. 58
	p2p-3	8, 717	31, 525	7. 23	5033	0. 58
Social Communication	UCIonline	1, 899	20, 296	21. 38	614	0. 32
Intra- organizational	Freemans-1	34	695	40. 88	1	0. 029
	Consulting	46	879	38. 22	2	0. 04

### 4.2 实验与分析

首先对这 14 个网络,应用 LB 模型计算出驱动节点数和驱动节点的分布情况.

图 3 给出网络中节点及驱动节点的度与平均度的比较分析数据,由图 3(a)可见,驱动节点避开了高度节点,且其驱动节点中低度节点的比例高于整体网络中低度节点所占的比例,图 3(b)给出了驱动节点中低于平均度的比例与网络中低于平均度的节点比例比值的曲线图,其值均大于 1. 因此,复杂网络控制中驱动节点的选取趋向选择低度节点不是一个随机现象.但是由于比值普遍较低(小于 2),该现象是否可以总结为普适规律,值得深入探究.

其次,分析这些实际网络引入免疫节点后的可控性.

对于随机免疫,在网络随机选取网络中 10%节点作为免疫节点,考察驱动节点增加比例的情况,并重复实验 20 次,取平均值.

对于目标免疫,分别按网络中节点的度指标 (DC)、介数指标 (BC)、紧密度指标 (CC) 的大小来定义免疫节点. 计算网络中所有节点每种指标值的大小,将网络中的节点按指标值的大小顺序划分为  $N_1, \dots, N_{10}$  这 10 个集合,保证每个集合中均具有 10%的节点,当节点数非 10 的倍数时,为保证  $N_{10}$  中的节点数量,在  $N_9$  中选择指标值低的节

点添加到  $N_{10}$  中.

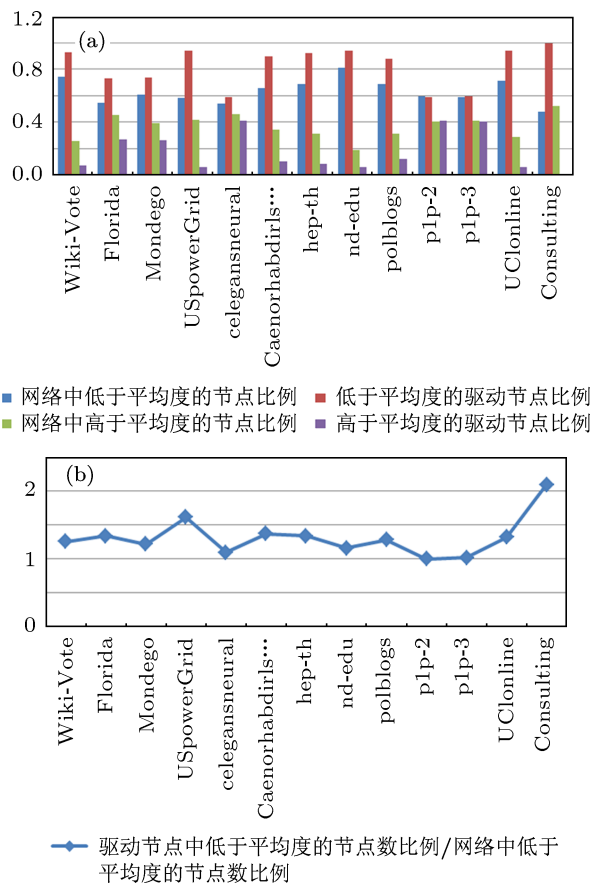


图 3 复杂网络中节点及驱动节点的度与平均度的比较分析图 (a) 网络中节点及驱动节点度分布; (b) 比值分析

图 4(a) 给出按随机免疫策略选择免疫节点时网络中驱动节点增加的情况,图 4(b), (c), (d) 分别给

出 DC, BC, CC 选择免疫节点时网络中驱动节点增加的情况, 图中虚线标出随机选择免疫节点所得驱动节点比例增加的平均值. 表 3 给出随机免疫策略

下各网络中驱动节点增加比例的相关统计数据, 及按这 3 个指标选取最重要的前 10% 的节点作为免疫节点时驱动节点增加比例的数据. 因为食物网、

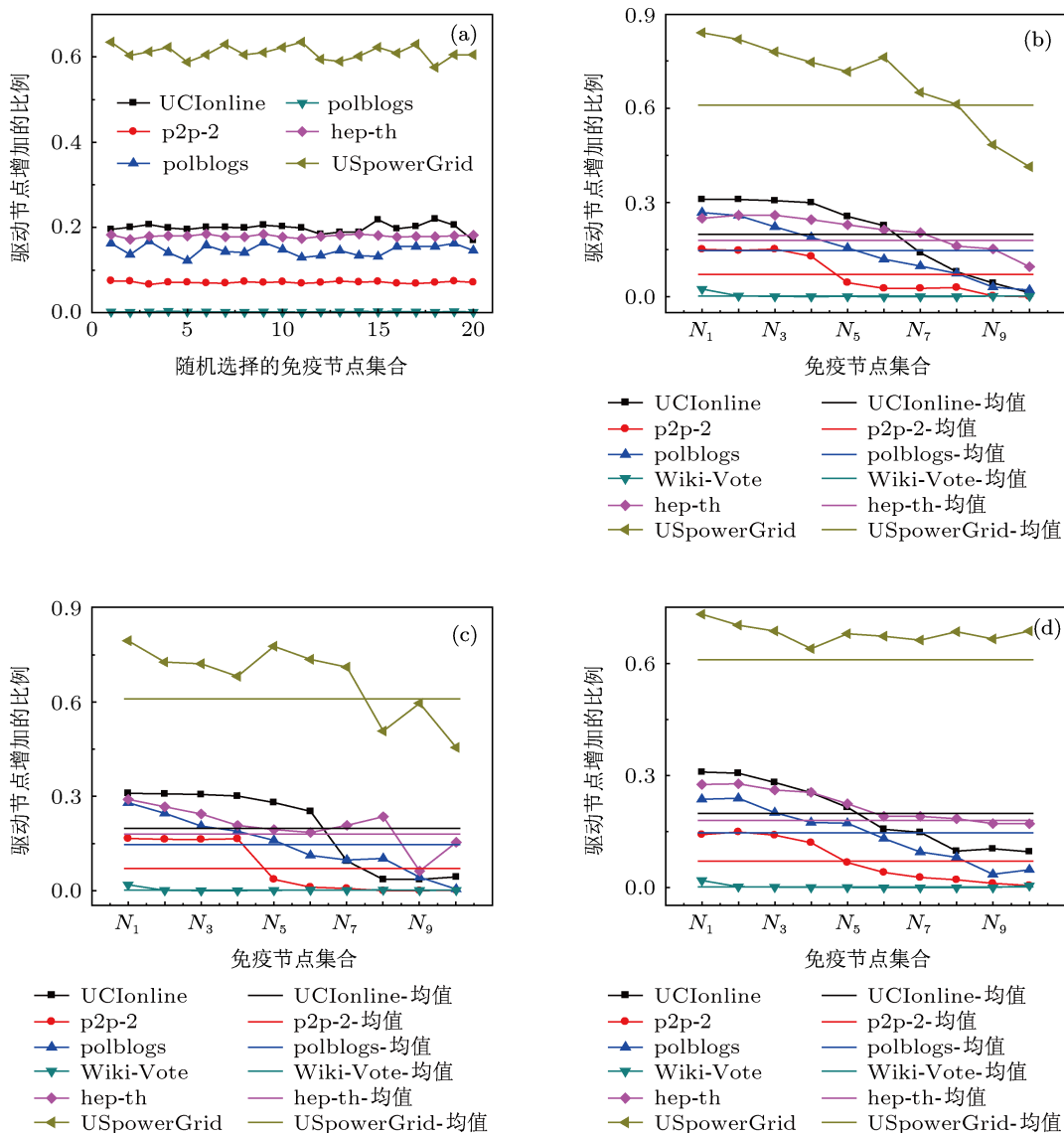


图 4 目标免疫时驱动节点增加比例 (a) 随机策略定义免疫节点; (b) 按度指标定义免疫节点; (c) 按介数指标定义免疫节点; (d) 按紧密度指标定义免疫节点

表 3 驱动节点增加情况的统计值

数据集	目标免疫			随机免疫			
	DC	BC	CC	平均值	方差	最大值	最小值
社会交际 -UCIonline	0.309558	0.30962	0.30962	0.198883	0.00012	0.219975	0.169483
因特网 -p2p-2	0.151438	0.166112	0.142047	0.071141	0.00000	0.074741	0.066132
万维网 -polblogs	0.267943	0.279904	0.236842	0.14677	0.00016	0.167464	0.12201
信任网 -Wiki-Vote	0.024704	0.019003	0.019215	0.00227	0.00000	0.003801	0.001267
引用网 -hep-th	0.249917	0.29029	0.276443	0.179771	0.00000	0.185018	0.172339
电力网 -USpowerGrid	0.84	0.794783	0.732174	0.610087	0.00025	0.634783	0.575652

神经网络、组织网规模比较小,在此不做分析,本文给出社会交际网、因特网、万维网、信任网、引用网及电力网的数据分析情况.

由图 4 和表 3 中数据分析得知,随免疫节点的重要度的增加,网络中驱动节点增加的比例增加,在目标免疫中将最重要的节点作为免疫节点时,驱动节点的增加比例大于随机免疫中驱动节点增加比例的最大值;当选择最不重要的节点作为免疫节点时,其驱动节点增加比例小于随机免疫中驱动节点增加比例的最小值;且当免疫节点的重要度位于前 50% 时,其驱动节点增加的比例均大于随机选取免疫节点时驱动节点增加比例的平均值.因此可得出结论,如果将网络中度数、介数和紧密度指标较高的节点作为免疫节点,将极大的提高控制复杂网络的难度.

这一结果表明:即使驱动节点倾向于低度节点,通过高度、高介数或高紧密度的节点阻断信号传播却能有效地影响网络的可控性.此结论与 Liu 等人<sup>[2]</sup>的结论中驱动节点避开高度节点的结论并不矛盾,而是对其的有益补充.

分析图 4 中数据,有以下两个现象:1) 信任网络 Wiki-Vote 将其最重要的 10% 节点作为免疫节点时,其控制难度增加明显,符合前述结论;而其余情况增加量不明显,其原因为网络中 66% 的节点为驱动节点,导致进一步提高驱动节点数量已经极为困难.2) 按紧密度指标选择电力网络 USpowerGrid 的免疫节点时,其驱动节点增加的比例均大于随机免疫的平均值,其原因为网络中节点的紧密度最大值仅为 0.08182,远小于万维网的 0.5185,且各节点紧密度的值分布相对均匀,因此用紧密度来衡量此网络中节点的重要度意义不大.因此,前述现象在进一步印证本文结论的同时,也说明复杂网络的领域性特点.

最后,从 LB 模型的核心理论基础 Lin 结构可控性理论,进一步分析此实验结果.

由于基于传播免疫的复杂网络可控性通过去除免疫节点的出边,使网络的拓扑结构发生了变化,因此分析此变化对网络中不可达节点与扩张的影响.

当将度数、介数或紧密度高的节点作为免疫

节点时,是否会导致其与邻域节点构成的子网络出现扩张,以及邻域子网络出现扩张的概率是否较高,这个问题的论证涉及到免疫节点的邻域节点之间的拓扑关系,较难通过理论证明.因此本文着重分析新增不可达节点的情况.不可达节点为不存在由输入节点到此节点的有向路径的节点.

表 4 度、介数、紧密度与出度的相关性

相关性	UCIonline	p2p-2	polblogs
度与出度	0.9711	0.7213	0.7717
介数与出度	0.8953	0.5336	0.5969
紧密度与出度	0.6130	0.7111	0.4582

很明显,在不考虑其他因素的前提下,去除出度大的节点的出边将直接导致新增的不可达节点的数目较多.从这个角度进一步分析节点的度、介数与紧密度与出度指标的相关性.表 4 基于三个网络,采用 pearson 相关系数分析三项指标与出度的相关性.其中 pearson 相关系数取值越趋近于 1,其正相关性越大.

可见,三个指标取值较大的节点作为免疫节点,去除的出边较多,将导致不可达节点出现的概率较大.因此,新增驱动节点数目较多,从而极大的提高控制复杂网络的难度.

## 5 结论

实际系统建模的网络中,并非所有节点“收到信号即转发”,针对这一问题,本文对 2011 年 Liu 等人<sup>[2]</sup>的复杂网络可控性理论进行扩展,提出基于传播免疫的复杂网络可控性理论,结果表明,如果将网络中度数、介数和紧密度指标较高的节点作为免疫节点,将极大的提高控制复杂网络的难度.

复杂网络系统可控性研究仍存在很多问题有待进一步研究分析,包括:免疫信号与扩张的理论关系,不同领域中控制的意义,输入信号与控制节点的连接选择问题,非线性时变复杂网络的控制等.

感谢东北大学信息科学与工程学院张锡哲教授的讨论.

- [1] Wang X F, Su H S 2008 *Advances in Mechanics* 751 (in Chinese) [汪小帆, 苏厚胜 2008 力学进展 751]
- [2] Liu Y Y, Slotine J J, Barabasi A L 2011 *Nature* **473** 167
- [3] Albert R, Jeong H, Barabasi AL 2000 *Nature* **406** 378
- [4] Kitsak M, Gallos L K, Havlin S, Liljeros F, Muchnik L, Stanley H E, Makse H A 2010 *Nature Physics* 888
- [5] Romualdo P S, Alessandro V 2011 *Phys. Rev. Lett.* **86** 3200
- [6] Centola D 2010 *Science* **329** 1194
- [7] Wang X F, Li X, Chen G R 2006 *Complex network theory and its application* (1st Ed.) (Beijing: Tsinghua University Press) p78 (in Chinese) [汪小帆, 李翔, 陈关荣 2006 复杂网络理论及其应用 (第一版) (北京: 清华大学出版社) 第 78 页]
- [8] Lin C T 1974 *IEEE Transactions on Automatic Control* **19** 201
- [9] Hopcroft J E, Karp R M 1973 *SWAT* 225
- [10] Cho A 2011 *Science* **332** 777
- [11] Magnus E 2011 *Nature* **473** 158
- [12] Xu D, Li X, Wang X F 2007 *Acta Phys. Sin.* **56** 3 (in Chinese) [许丹, 李翔, 汪小帆 2007 物理学报 **56** 3]
- [13] Wang Y Q, Jiang G P 2011 *Acta Phys. Sin.* **60** 060202 (in Chinese) [王亚奇, 蒋国平 2011 物理学报 **60** 060202]
- [14] Wasserman S, Faust K 1994 *Social Network Analysis: Methods and Application* (New York: Cambridge University Press) 65–94
- [15] Freeman L C 1977 *Sociometry* **40** 35
- [16] Costenbader E, Valente T W 2003 *Social Network* **25** 238

## Controllability of complex networks based on propagation immunization\*

Lü Tian-Yang<sup>1)2)†</sup> Piao Xiu-Feng<sup>1)</sup> Xie Wen-Yan<sup>1)</sup> Huang Shao-Bin<sup>1)</sup>

1) (College of Computer Science and Technology, Harbin Engineering University, Harbin 150001, China)

2) (College of Computer Science and Technology, Tsinghua University, Beijing 100084, China)

(Received 16 November 2011; revised manuscript received 20 February 2012)

### Abstract

Control of complex network reflects humans' comprehension of complex system and the ability to reform it. Up-to-date research establishes the controllability theory of the complex networks based on linear system control theory. The theory could find a minimal set of driver nodes which controls all nodes' state in a linear time invariant complex network with any topology. However, this theory does not take into account the immune node or failure node which blocks the control signal. While inheriting the advantages of the theory, in the paper we first refine the complex network control model based on propagation immunization. Second we adopt four methods which belong to random immunization strategy and targeted immunization strategy to determine the immune nodes, and analyze the controllability of 14 real networks. The experimental results show that when the nodes which have higher degrees, betweenness or closeness are treated as immune nodes, the control of complex networks will become more difficult.

**Keywords:** control of complex network, propagation immunization, immune nodes

**PACS:** 05.65.+b, 02.30.Yy

\* Project supported by the National Natural Science Foundation of China (Grant Nos. 60903080, 60093009), the National Key Technology Research and Development Program of the Ministry of Science and Technology of China (Grant Nos. 2009BAH42B02, 2012BAH08B02), the Fundamental Research Funds for the Central Universities (Grant Nos. HEUCF100603, HEUCFZ1212, HEUCFT1206), and the Scientific Research Fund of Heilongjiang Provincial Education Department (Grant No. 12513050).

† E-mail: raynor1979@163.com